Two and Four Year Olds' Understanding of Space: A Comparison of Imitating and Describing Directed Motion Events

Honors Research Thesis

Presented in Partial Fulfillment of the Requirements for Graduation
"*with honors research distinction in Linguistics*" in the Undergraduate Colleges of
The Ohio State University

by

Hartman Brawley

The Ohio State University
March, 2012

Project Advisor:
Professor Laura Wagner, Department of Psychology

ABSTRACT

In this study we investigated the relationship between language and thought and whether it changes with language experience over time. As language skill matures in children, does their linguistic representation of an event influence their non-linguistic representation of the same event? We also investigated the nature of this influence; that is, does the linguistic perspective globally inform all non-linguistic tasks? Does it locally inform those nearby in time? Two-year-olds (mean age 31.91 mos) and four-year-olds (mean age 53.17 mos) were presented with Directed Motion (DM) events and analyzed them in a linguistic and a non-linguistic domain. DM events consist of a body in motion moving between points in space. The motion contains a trajectory, the Path, and the characteristic quality of the motion, the Manner. The linguistic task was elicited description, and the non-linguistic task was the forced imitation choice with two phases. In the first phase (Direct) children exactly imitated a DM event performed with a stuffed animal on a foam core ramp. In the second phase (Choice) children imitated modeled events on an environment that differed from the experimenter's in such a way that they could only imitate either the Path or the Goal of the modeled event. The study found that two-year-olds did not prefer components in either task, but showed an order of task effect. When they imitated first, they directly imitated all components correctly more often and described Path more often. Four-year-olds did not show this order of task effect. This suggests that for younger children, the strategy used for whatever task they performed first carried over into the subsequent task. Older children did not perseverate, having already solidified the means with which they carry out tasks in different domains. Also, this suggests that gaining language experience changes the relationship between language and thought over time.

ACKNOWLEDGEMENTS

This project owes its completion to the support and care of many. First I would like to thank my project advisor, Dr. Laura Wagner, not just for her guidance and instruction, but also for all of her encouragement along the way. My committee members, Dr. Shari Speer and Dr. John Grinstead, also deserve many thanks for their help and suggestions with the project. I would also like to thank the Departments of Linguistics and Psychology, along with the support from Social and Behavioral Sciences and the Buckeye Language Network, who provided funding for the project, and members of the Developmental Language and Cognition lab, who assisted with various aspects of the project. Finally, I would like to thank Christine Szostak for listening and helping me talk about my work, my loving parents and sister for all their moral support, and Monte Main for keeping me sane, motivated, and caffeinated.

TABLE OF CONTENTS

Introduction

Although one of the goals of language use is to organize observations for the efficient presentation of information, the information that ends up getting conveyed does not perfectly reflect the observation. Something as simple as a ball dropping onto the floor can be described in infinitely many ways, ranging from terse - "The ball dropped to the floor." - to elaborate - "The small red ball rolled off of the table and fell at a rate of three feet per second to the tile floor, whereupon it bounced." Though the speaker can go into great length and detail describing the event, especially if prompted to, there will always be aspects of the event that were noticed but chosen not to be included in the description (e.g. in the latter description I mentioned that the floor was tile but not that the table was wooden). Furthermore, there is no perfect characterization of an event. Events as they exist in the world are made up of an infinite number of factors, many of which could be talked about. There are factors that exist outside the scope of human perception, such as the molecular composition of the ball, contextual factors that may not be relevant to mention or have already been mentioned, such as the fact that the ball belongs to me and it was a gift, and spatial and temporal boundaries that define the event, such as the fact that the ball fell today at 3:01 PM and indoors. Because there are no intrinsic aspects of events that are obligatory to mention, it is up to the viewer, and more specifically the mapping of event representations onto syntactic forms in the viewer's native language, to impose such architecture upon the event. These syntactic forms are finite in number, and since languages are set up differently to express events, speakers are encouraged talk and possibly think about events in such a way that reflects their language's structure.

Different languages can create different pictures of the same events. For example, consider the following event: a boy climbs a tree and rests at the top. In English, a speaker is

likely to describe trajectory of the motion as in (1), while in Spanish a speaker is more likely to describe the endstate of the motion as in (2), from (Slobin 1996).

(1) The boy climbed the tree.

(2) El niño está subido en el árbol.
(The boy is climb-PART in/on the tree.)
"The boy climbed the tree."

What accounts for these differences? We can assume that the two speakers experience the event similarly, in that they both have human sensory organs. Moreover, infants and young children without much experience with the language will have similar mental representations of the event. But when speakers mature, will the accumulated habits of using their language change the way they read the input, so that certain pieces of the event are more salient and others ignored, uniquely from speakers of other languages? That is, does language learning affect how we think about the world?

This thesis investigates these questions. First we explain Direct Motion events (henceforth DM events) and their value to the study of language and thought. We then provide an overview of the leading theories about the relationship between language and thought and the Whorfian linguistic determinism hypothesis. Then, we discuss difficulties inherent in studying language and thought, some potential methodological solutions, and finally introduce the current study. From there we present our Methods, Results, and Discussion.

Directed Motion Events

As previously discussed, languages differ systematically in their representations of events. Talmy (1985) described two major typological categories for languages with different rhetorical representations of space, and more specifically of DM events. DM events are defined by a Figure, something which undergoes motion, moving between Grounds, points of reference

in space. A Ground may be the location where the Figure begins moving (these are known as Sources), where the Figure stops moving (Goals), or simply locations past which the Figure moves. The motion itself can be characterized along two different dimensions, Manner and Path. Manner refers to the idiosyncratic character of the motion (e.g. *walking* vs. *jumping* in English, verbs describing specific actions). Path is the directionality of the motion (e.g. *entering* vs. *exiting* in English, note that these do not describe a specific action). There also exist verbs that describe both (though arguably not in English).

Returning to typology, languages are categorized according to where Path information is encoded in the utterance. Verb-Framed languages, such as Spanish, Greek, and Japanese, encode Path in the main verb and Manner in optional constructions such as prepositional phrases, adjuncts, or secondary verbs. Satellite-Framed languages, such as German and English, encode Manner in the main verb and Path optionally. Consider the following example sentence in English, a Satellite-Framed language:

(3) The boy ran from the house to the street.

In this example, the Figure of the DM event is *the boy*. Two Grounds are mentioned, a Source and a Goal. The Source is encoded by *from the house*, where the boy began moving, and the Goal is *to the street*, where the boy finished moving. The Manner of the motion is encoded by the action verb *ran*, which is the main verb. The Path is implied by mention of the Source and Goal, but no trajectory is explicitly mentioned. Now consider this same sentence in Japanese, a Verb-Framed language:

(4) Otokonoko-wa uchi kara michi made hashitte itta.
(boy-GEN house from street to running went)
"The boy ran from the house to the street."

In this sentence, the same elements are represented. The Figure, *otoko no ko*, "the boy," the Source, *uchi kara*, "from the house," and the Goal, *michi made*, "to the street," are all expressed similarly to the English example, but the Manner and the Path are expressed differently. Rather than in the main verb, the Manner is expressed by the gerund form *hashitte*, "running." The Path is expressed by the main verb *itta*, "went." *Itta* and other forms of the Japanese verb meaning "to go" express information about the trajectory of the motion relative to the position of the speaker, whereas forms of the verb *go* in English do not. Over time, typological differences between languages contribute to preferences for event components in their rhetorical styles. In other words, there are language-specific patterns determining what event components speakers encode in their descriptions. If these preferences are also present in their non-linguistic intuitions about events, it would signal a linguistic influence on their non-linguistic thought.

Non-linguistic cognition also has an influence on linguistic descriptions. For example, in their 2005 study, Lakusta and Landau showed typically developing children, children with William's Syndrome, and typically developing adults 34 movies of DM events and asked them to describe them. In all types of movies – change of possession events, change of state events, and attachment/detachment events, both the Source and the Goal were salient, and all participants were much more likely to mention Goal Paths and omit Source Paths. There may also be pre-linguistic tendencies that children have as they approach learning how to use language to describe DM events. Wagner and Lakusta (2009) have argued that pre-linguistic infants have event representations that make use of the same information as their caregiver's language's semantic categories of DM events as described by Talmy (1985) (Manner, Path, Goal, etc.). Choi and Bowerman (1991) provided further evidence of this. They recorded the

speech of two English-speaking and two Korean-speaking children longitudinally from the age of about 1 to 2 years and discovered that English and Korean children encode Path in their event descriptions differently, and according to the lexicalization patterns of their respective languages as described by Talmy's typology. The English speaking children used Path particles to encode Path in a variety of events, whereas Korean children used Path verbs first with caused, transitive motion, and learn verbs for intransitive and spontaneous motion later. Choi and Bowerman concluded that these patterns, which align with those shown by adult speakers of English and Korean, suggest that children attend to language-specific constraints on representing space well before they learn how to encode space in language.

Note that in either language the typological of encoding information are not completely rigid. That speakers become accustomed to the descriptive conventions of their language does not mean that the conventions are obligatory. Natural descriptions of events can be uttered that do not conform to typological expectations. Consider (7), one such sentence in English:

(7) Sam entered the room.

Being a Satellite-Framed language, English is expected to express Path information outside of the main verb. In this case, the verb "entered" defies that expectation by describing the trajectory of Sam with respect to the room – first he was outside the room and then he was inside the room. These descriptions are generally less frequent, require greater syntactic marking, and are longer than their typologically typical counterparts.

The Whorfian Hypothesis

The Sapir-Whorf linguistic determinism hypothesis attempts to explain the relationship between language and thought by making the strong claim that native speakers of two different languages are compelled to attend only to those dimensions of their experience that suit the

syntactic categories of their language and ignore those that do not. A weaker version of the hypothesis, sometimes referred to as Neo-Whorfian linguistic relativism, suggests that the effect is less rigid, facilitated both by linguistic experience and cognitive constraints on experiencing the event. A corollary of this theory is that cognition is affected gradually over time as more language experience is gained. According to this interpretation, all speakers begin with a baseline method for analyzing of events that is unaffected by language-specific perspectives. As a speaker gains experience with his or her native language, over time this baseline method will be blended with gradual interference from the language's constrained perspective of the world, in space and motion as Talmy describes, but also in all other domains of thought. This hypothesis predicts that speakers will increasingly draw upon this linguistic method even in non-linguistic situations (Slobin, 1996, 2003, 2004, 2005, 2006). It also predicts that speakers will make use of this method to fulfill the demands of any task, and regardless of the demands of any recent tasks. For this reason we will refer to this perspective as the Global relationship between language and thought.

There is another possible explanation of the relationship between language and thought, described by Gennari et al. (2002) as the Language as Strategy Hypothesis. It suggests that speakers can rely on the way they prepare to use language to help them perform certain non-linguistic thought tasks that rely on similar demands as language. For example, Lucy (1992) demonstrated that speakers rely on their particular language's existing semantic categories to categorize novel objects. Adult speakers of English and Yucatec Maya were presented with 17 triads of objects, one original and two alternates: one that different in its shape and one that differed in its material composition. An example triad would be ceramic bowl (original), a ceramic plate (shape alternate), and a metal bowl (material alternate). Participants were asked to

decide which alternate was more similar to the original object. English speakers reliably chose the material alternate, and the Yucatec Maya speakers chose the shape alternate. Lucy concluded that because English forms object contrasts on the basis of shape more than material composition, and Yucatec Maya material more than shape, that these linguistic habits drove participants to find distinct salient differences among the same objects. But only in executing certain tasks can the speaker benefit from relying on their linguistic perspective. Papafragou & Selimis (2010) asked Greek and English speaking participants to view 48 motion events, describe them, and then match each one to one of two other events, each of which varied by either the Manner or the Path. Participants categorized them according to verb lexicalization patterns in their native language - English speakers grouped events with similar Manners and Greek speakers with similar Paths. However in a second experiment participants were not asked to describe the events and any linguistic cueing was removed from the procedure. In that case the effect did not appear. Importantly, this finding demonstrates the effect of the language in the instructions and descriptions contributing on performance in a subsequent, non-linguistic task, but not a Global effect from language onto a non-linguistic task even without direct linguistic cues. This strategy formation based on the presence of linguistic cueing we will refer to as the Local relationship between language and thought.

Global and Local effects from language onto thought do not have to be exclusive of one another. It is possible that for some tasks mature speakers' interpretations of events are always affected by their language (Global), but for other tasks, a recent, direct appeal to linguistic knowledge is needed for the linguistic influence to appear (Local). Gennari et al. (2002) demonstrated this by presenting English and Spanish adults with 36 videotapes of DM events. Participants were in one of three conditions, one where they had to describe each event, one

where they did not, and one where they instead were asked to repeat nonsense syllables after each event. Then, all participants were shown 108 DM events, 36 of which they were shown previously, and were asked to indicate for each event whether they had seen the event before. Lastly, they provided similarity judgments on the same set of events. They found that speakers of either language that described beforehand did not recognize events better according to the information they encoded in their descriptions. However, they did judge events with the same information encoded in their descriptions as similar more often. This suggests that some non-linguistic tasks may access to language all the time, whereas others do not unless something triggers a linguistic analysis of the event first.

Experimental Concerns

There are certain difficulties inherent in studying language and thought. One of these is studying language in isolation of other thought processes. Gleitman et al. (2007) showed in an eye tracking study that event apprehension and processes of linguistic planning do not operate serially, but cooperatively and in unison. They presented participants with a fixation cross in the center of a screen followed by black box for only 60-75 milliseconds on either the left or right side of the screen followed by a scene involving two subjects, one on either side, that they described. The box directed their gaze to one of the two figures at onset of the scene, therefore making it more salient at first. The scenes consisted of an action carried out by an agent onto a patient (for example, a dog chasing a man). The prediction was that if language and scene apprehension happen separately, then participants should describe the scenes according to the highest frequency patterns, such as active voice over passive voice, animates mentioned before inanimate objects, etc (in the example case, "a dog chased a man" is higher frequency than "a man was chased by a dog"). Even the though participants reported not noticing the black box, it

successfully guided them to describe the scene beginning with the figure in the location that the black box had appeared in just before the trial. Furthermore, participants described the figure to which their attention was directed even when it resulted in the lower frequency description ("a man was chased by a dog"). This suggests that participants were gaining an initial "gist" of the scene at the same time they were formulating their description, both within 200 milliseconds of viewing the event.

Another issue is that processes of language must make use of non-linguistic cognition (e.g. a gist must be obtained first to form a description). Participants unconsciously use linguistic knowledge to perform a task, which makes it difficult ascertain whether an effect from language onto thought is Global or Local. A careful manipulation is therefore needed: on the one hand, linguistic interference must be minimized in a non-linguistic task to find a Global effect, and on the other hand it must be clearly telegraphed if it is introduced to evoke a Local effect.

Experimental Solutions

One common way to test this effect from language onto thought is comparing the non-linguistic intuitions of adult speakers of Verb-Framed and Satellite-Framed languages. If Global linguistic relativism holds true, mature speakers that have already formed their habits of expression will think non-linguistically different about events from speakers of other languages. Work analyzing adults has generated somewhat split findings. Papafragou et al. (2002) presented two experiments, one asking participants to describe pictures of DM events, and another asking different participants to perform a memory task in which they are shown pictures of DM events and later asked if they recognize a new set of events. The Greek and English speakers performed according to typological expectations in describing the DM events, but this

did not predict memory performance, which was similar between English and Greek speakers. This is evidence against the relativist claim that there is a Global effect from language onto thought. On the other hand, Billman and Krych (1998) do claim to find an effect from language onto later recognition memory of events. They showed participants a video of 24 events accompanied by spoken verbs and then the next day showed them 24 movies, 8 of which were the previously viewed targets and 8 foil events that contained either a different Path or Manner, and asked them how certain they were that they had seen the event before. Participants were more correct in distinguishing the event if it was a foil of the same information encoded by the verb than when it was a foil of different information. In the control condition in which no language accompanied the presentation of the event, participants were equally as accurate at distinguishing foil events of either kind from model events. This demonstrated that when linguistic information was provided when viewing the event, it was stored as part of the event representation along with the visual information. The Local perspective can account somewhat for this finding. Recall that the effect of language on thought is Local and sensitive to priming, so if a direct appeal is made to linguistic knowledge when initially analyzing the event, it will be used to make subsequent non-linguistic analyses about the event. All the same, the results leave the debate about linguistic relativism an open question.

Another possible analysis is that of comparing younger children and older children within the same language. Having more language experience, older children should show stronger Global and Local effects of language than younger children who have less language experience. This is useful because of children's understanding of space and encoding of space. Understanding space is a fundamental cognitive faculty that is available to infants (Papafragou & Selimis, 2010). Language is not. Although terms describing space (such as "up" and "down" in

English) are learned very early on cross-linguistically along with some notions of language-specific spatial semantic categories, it takes more experience to ultimately master typical language patterns of mature speakers. Therefore, younger speakers may have non-linguistic expectations about space non-linguistically, but it will take time for linguistic structural patterns to change these expectations. Moreover, if change does occur over time, younger speakers should perform differently than older speakers. There is some evidence to suggest that as language skill increases conformance to typological norms of expression does as well. Papafragou et al. (2006) showed that young speakers of English and Greek, in a task where they described events, increasingly aligned to the lexicalization patterns of their respective languages with age. Choi and Bowerman (1991) found in their broader study of recording a Korean child's and an English child's language from the age of 1-2 years that right from when children began to speak they were sensitive to adult speakers' means of describing space and described it differently from their language counterparts. The English speaking child used prepositions to describe Path from very early on, whereas the Korean child used a variety of Path verbs. On the other hand, Slobin (1996) argued through his study of children of many different languages describing Mercer Mayer's picture book *Frog, where are you?* (which does not contain text, only illustrations of events) that although children may show some features of mature language use from when they begin speaking, it is not until around the age of three years that they master the rhetorical style of their language. Rhetorical style in this context refers to the dimensions of space that are either expressed, made implicit, or omitted by typical language use of a certain language group. The style is learned partly by what is obligatorily encoded by verbs in the language according to Talmy's typology, but also culturally from the accumulated input of how mature speakers describe events (for example, adults around me always mentioned trajectories in

their event descriptions, so now unconsciously so do I). As older children gain more experience with their language's rhetorical style, they should show stronger Global and Local effects of language, if they exist.

As previously discussed, it can be difficult to find a non-linguistic task that is free of linguistic influence. A unique alternative to previously used measures is the imitation choice paradigm used by Wagner et al (2008). They presented 34 two-year-olds with DM events acted out using a stuffed animal on a foam board. The board consisted of two level planes connected by a ramp, and a differently colored plastic bowl on each level, one face-up and the other face-down to appear distinguishable. The modeled DM events consisted of one of two possible Manners, hopping the animal or sliding the animal, one of two possible Paths, going up the ramp or down the ramp, and one of two possible Goals, the bowl on top or on bottom. The Goal objects never moved, and the modeled events always respected the flow of the Path and Goal (i.e. an "up" Path always led to the goal on top and vice versa), so there were four possible event configurations. In the first, Direct phase of the task, participants were modeled each of the four possible DM events and after each one were asked to imitate the event on the same board. In the second, Choice phase, participants were instructed to carry out their imitations on a separate board given to them, which differed from the experimenter's in that the Goal objects switched positions. The events were again modeled to the participants and they were asked to repeat the event on their board, forcing them to make a choice: they could either imitate the modeled Path at the expense of the Goal, or imitate the modeled Goal at the expense of the Path. Overall, the children showed a significant bias for Path over Goal in their choice imitations. This paradigm is useful as a non-linguistic measure because it breaks ground by moving away from previously used non-linguistic measures such as memory recognition and categorization, in case these

methodologies somehow call upon linguistic processes in participants (say, if memory is partly stored using language or categorization always accesses lexical processing). Planning for imitation should not access any linguistic strategy.

Present Study

The current study examined two-year-olds and four-year-olds analyzing the same DM events in a linguistic and a non-linguistic task. Based on Slobin's assertion that mastery of rhetorical style happens somewhere around the age of 3 years, there should be adequate linguistic development between our two age groups. The linguistic task was elicited description, and the non-linguistic task was the imitation choice paradigm used by Wagner et al (2008). Past work has used static scenes of DM events for description stimuli (Papafragou et al., 2002; Gleitman et al., 2007), but the current study used movie clips to preserve the natural dynamicity of natural motion events, and to make them more parallel to the modeled events in the imitation task. Participants also performed both tasks using the same events, providing a direct, within subjects comparison. The order of tasks was counterbalanced, so if a Local effect exists it should appear for participants in the description first condition.

Predictions

Based on previous work, our assumptions about how DM events are structured, and the capabilities of children in each age group, we have some predictions about children's performance in this study. Older children should have longer descriptions, include more components in their descriptions, and perform better in Direct imitation simply because of their higher cognitive skills over younger children. With respect to the existence of Global and Local effects, several hypotheses can be made:

Hypothesis I: If there is a Global effect but not a Local effect from language onto thought, then older children should prefer different information than younger children for both task orders. Older children should show a stronger effect because greater language experience should bring to bear a greater linguistic influence onto their thought.

Hypothesis II: If there is a Local but not a Global effect, then older children should prefer different information than younger children, but only in the description first condition. Since Local effects depend on the prominence of the linguistic perspective when carrying out a non-linguistic task, the description first condition only should provide this prominence.

Hypothesis III: If there are both Global and Local effects, we expect the same outcome as Hypothesis I, but the effect should be stronger in the description first condition because of the linguistic priming.

Hypothesis IV: If there are neither Global nor Local effects, than the two age groups should prefer the same information in both tasks, and regardless of task order.

Methods

Participants

15 two-year-old children and 16 four-year-old children participated in this study, 17 of which were boys and 14 of which were girls. There were 7 girls and 7 boys in the imitation first condition, and 8 girls and 9 boys in the description first condition. The top of the two-year-old age range was 35.57 mos, the bottom was 27.8 mos, and the mean was 31.91 mos. The top of the four-year-old range was 58.6 mos, the bottom was 45.27 mos, and the mean was 53.17 mos. Participants were all recruited and run at a satellite research space at the Little Kidspace© at the Center for Science and Industry (COSI) children's museum in Columbus, Ohio. The experiment

lasted about three minutes, during which the parents almost always chose to be present, and afterward the children were offered hand stamps for their participation.

Another 61 children participated but were not included in the study. 25 of these refused to participant beyond the very initial stages of the experiment (i.e., being introduced to the stimuli), another 2 only completed less than half of the imitation task trials, 25 completed less than half of the description task trials, and all but one of those were two-year-olds, a testament to how difficult the description task was for the younger participants. Another two two-year old participants were removed because all of their descriptions were unintelligible and were therefore unable to be coded. Seven participants were removed due to experimenter error.

Stimuli

For the imitation task DM motion events were presented by moving a Cookie Monster stuffed toy on foam core ramps that were roughly 12 inches long, 11 inches wide, and 9 inches high. The boards contained a lower level and an upper level connected by a straight ramp, providing two possible Paths – up or down – and one Goal object on both the upper or lower level where the motion would terminate. During the Choice phase of the imitation task the experimenter would model the event on a board with one Goal object orientation, and the participant would have to imitate it on a separate board on which the Goal objects were in the opposite locations. The Goal objects were either an orange baby food bowl turned upside-down or a red bowl right side up (see Figure 1). A board was also created for the Direct phase of the imitation task, which had the same dimensions of the other boards (see Figure 2).

Figure 1: Direct imitation environment and Figure

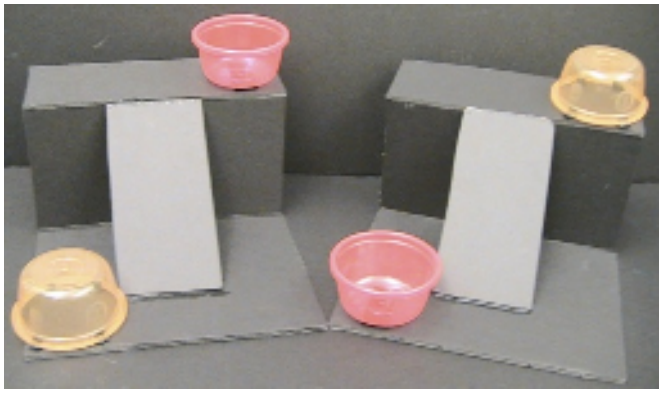

Figure 2: Choice imitation and description task environments

For the description task a woman from the lab was recorded moving a Curious George stuffed animal on one foam environment that was identical to one of the Choice phase foam boards, each clip being a DM event modeled by the experimenter in the Choice phase of the imitation task. These videos preserved the audio of the animal making contact with the board but no speech or other sounds. The clips were cut in Final Cut Studio.

For both tasks, the DM events displayed were controlled so that they all presented the same number of event components, but between movies the content of the components varied. There were two possibilities for each component - the Manner of motion was either hopping or sliding, the Path the animal travelled was either up or down the ramp, and the Goal of the motion was one of the two bowl objects. However, the direction of the Path would also determine the Goal of the event so that the trajectory always ended at the expected destination, for example the animal would always move up the ramp to end up at the top bowl and vice versa. Combining these possibilities created 4 events, all of which were both modeled to participants in the imitation task and shown as movie clips in the description task.

Procedure

Participants all performed the description and the imitation task, and were balanced so that half performed the linguistic task first and the other performed the non-linguistic task first. This way it was possible to have a within subjects analysis of children's linguistic and non-linguistic responses to the same events, and also see if there was a local effect from language onto thought in either age group.

Description Task:

The linguistic task was a simple description task, in which children were shown 4 clips of DM events, and after each clip were asked to describe what they saw. Specifically, children were asked, "What happened?" "Tell me what that movie was about," or "What did you see?" Questions did not include content pertaining to the agent, animal, motion, or foam environment to avoid cueing children on what to describe. Children were allowed to see the movie three times total if they refused to describe the clip after viewing it. After three viewings the experimenter moved on to a different clip and would not return to it. In order for their data to be included in the data analysis, children had to provide a description for at least two of the four clips. In a further effort to reduce noise from younger participants, a further caveat was introduced. Younger children also had to provide at least one description that included event components. This eliminated two younger participants whose descriptions were very short such as, “Monkey,” and did not provide content about the event components.

Imitation Task - Direct Phase

The non-linguistic task was the forced choice imitation task used by Wagner et al. (2008). First, in the Direct phase, the experimenter modeled two DM events on the practice board and after each asked participants to repeat it on the same board. Each time the participant was asked

to perform, the experimenter would only provide simple instructions that would not include any event components, draw attention to any portion of the board, or mention the stuffed animal. Examples of the instructions include: “Now it’s your turn to play,” “Can you play the same game?” “Play that game for me now,” etc. The experimenter then presented the same two events once more. Sometimes participants, especially the younger ones, would engage in creative play in the Choice phase in which they enacted components not modeled by the experimenter or different events altogether, such as playing under the ramp or off of the playground. Such creative play was considered uncodable. The Direct phase provided a non-linguistic inclusion measure that was parallel to that of the description task, in that neither Direct imitation or description methodologically pressured participants to encode particular components.

Imitation Task - Choice Phase

In the Choice phase, children were given a board of their own to imitate the events on that was different from the experimenter's, in that the goal objects were in opposite locations. In this way, the children must choose between imitating the modeled Goal correctly at the expense of the Path, or vice versa. The experimenter modeled four DM events on his own board, after each of which he asked participants to repeat it on their board. Some children, especially four-year-olds, looked incredulous, or tried to move the positions of the Goal objects to match the experimenter's board, or stated "I can't do it." In these few cases they were assured that they could and were instructed to give it their best try. All children who raised an objection of this nature still completed the task. Children had to complete at least 2 of the four Choice trails along with 2 description trials to be included in the analysis.

Coding

Description Task: For analysis the descriptions were broken down into clauses. Each clause was coded for inclusion of Path, Manner, and Goal event components. A clause for the purposes of this study comprised a main verb and all of its arguments and complements. In other words, every time a new main verb was encountered it was considered a new clause of the description. A sentence fragment was coded as one clause, and repetitions or self-corrections of clauses were not coded. Breaking the descriptions into clauses provided a better comparative measure of linguistic prowess between two and four-year-olds, and a way to see when in their descriptions they encoded specific information. To this end, average clauses per description, number of components per description, and number of components per clause were also recorded for all participants. Inclusion was coded by a score of 1 or 0 for each component in each clause, and by what syntactic feature it was included. Explanations for each event component follow.

In English, Manner is usually encoded in the main verb as in *hop* or *slide*, coded in the current study as specific-verb, but may also be encoded by adjuncts such as *by hopping*, coded as adjunct. Not all verbs that describe motion describe Manner of motion. Examples of these are *go*, *got into*, or *moved*. Clauses containing such verbs were coded as non-specific. If no motion verb was used, or the clause described something outside of the motion event, it was coded as omit.

Path of motion was considered included if directional movement on the ramp was mentioned. This was possible using very high frequency particles such as *up* and *down*, coded prep only, but could also be expressed by using Path verbs such as *climb* and *fall*, coded verb only. Path verbs received credit for Path inclusion if they were used to describe movement on the ramp. Combinations of verbs and prepositions, such as *fell down*, were coded verb prep. *Up*

and *down* were also used in ways that did not describe the directed motion, such as *sat down* (51.73 mos), which describes the Goal of the event but not the directional movement. Clauses containing these constructions would not receive credit for Path inclusion. These clauses, along with clauses containing no mention of Path, were coded omit.

For the Goal of the event, mention of the Goal objects or what the animal did at the Goal object earned credit for Goal inclusion. This was achieved by using a prepositional phrase such as *to the*..., or mentioning an action at the Goal object as in the Path example *sat down*, where it is implied that the animal sat on the Goal object. Clauses containing information about the physical Goal location were coded physical clause, and clauses containing more than just Goal information were coded physical. Clauses without Goal information were coded omit.

Note that it is possible for participants to form a description of the experimental events including all components, "He hopped up and into the bowl," or none of them, "Look at him go!". Inclusion scores for the three event components for each clause were averaged per subject over all of their descriptions, which had to be at least two per subject. In this way each subject's score represented how characteristic it was for that subject to provide a description for an event that included a certain component.

Imitation Task - Direct Phase

Direct imitation trials were coded for correct imitation of Manner, Path, and Goal event components. For each trial, if the Manner was correctly imitated from the modeled event, it received a Manner inclusion score of 1. Imitating multiple Manners, the incorrect Manner, or no Manner earned a score of 0. These criteria were also used for Path and Goal. These inclusion scores were averaged for all trials per subject.

Imitation Task - Choice Phase

For Choice trials a Path Bias score was given for the imitations. If the participant imitated the correct Path at the expense of the modeled Goal, they were given a score of 1 for that trial. If they imitated the correct Goal at the expense of the Path, they earned a score of -1 for that trial. It was also possible to show no bias in imitation by either correctly imitating both Path and Goal (e.g. sliding the animal down the ramp and putting it in the top bowl, also breaking the flow of the event), or correctly imitating neither of the target components. Path bias scores were averaged for all trials per subject. Though Manner was not included in the Path bias score, for each trial a Manner inclusion score was also recorded.

Results

Imitation Task – Direct Phase

A repeated measures ANOVA was conducted with task order (imitation first vs. description first) and age group (two-year-olds vs. four-year-olds) as between subjects independent variables and event component (Path vs. Manner vs. Goal) as a within subjects independent variable. The dependent variable was the proportion of correctly imitated elements, calculated from the average component inclusion scores from each participant's imitations. The analysis found a main effect of age ($F(1,27) = 12.146$, $p = .002$) – four-year-olds correctly imitated more elements than twos – and task order ($F(1,27) = 5.234$, $p = .03$), and an interaction ($F(1,27) = 10.362$, $p = .003$) between age group and task order. The task order main effect and the interaction both stem from the fact that two-year-olds perform better overall in the imitation first condition, and four-year-olds are at ceiling in both conditions. There was no main effect of component inclusion or interactions with component inclusion – components were included at

equal rates in both conditions. This is not surprising since in this phase the children could and were encouraged to imitate all event components. Figure 3 shows the Direct imitation inclusion scores for both age groups split across conditions, showing no preference for any event component but a drop in overall performance for two-year-olds in the description first condition.

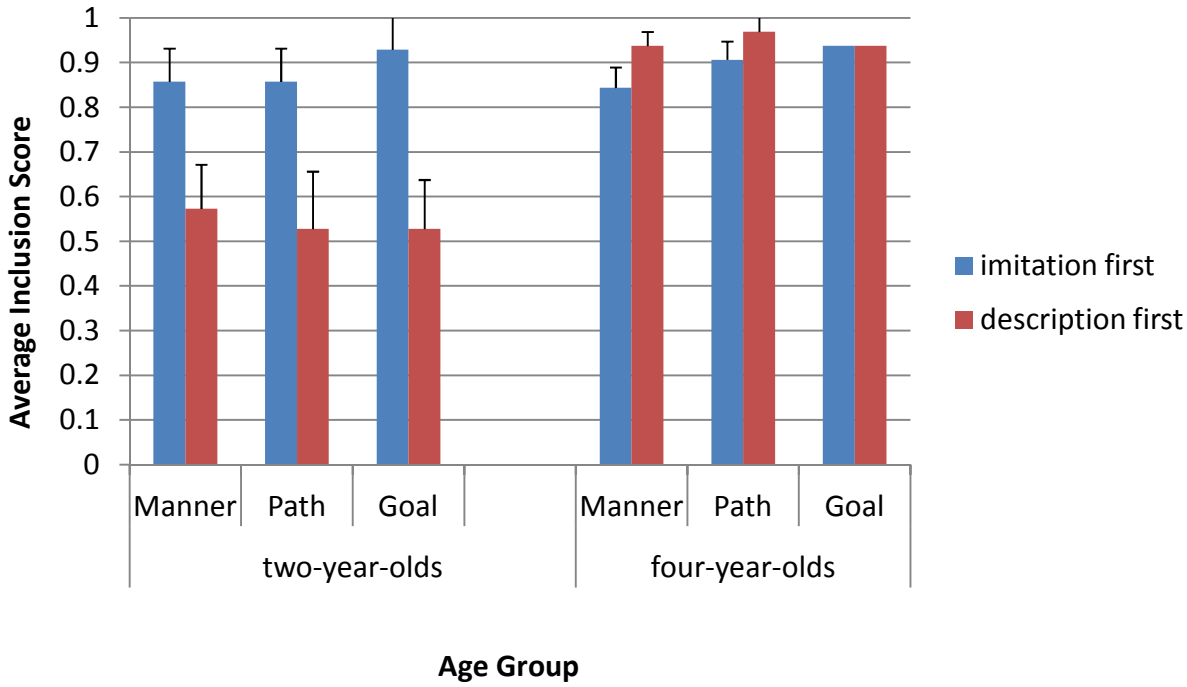

Figure 3: Inclusion in Direct imitation by each age group, broken down by task order condition.

Imitation Task – Choice Phase

In the Choice phase, recall that the stimuli events were set up so that participants were discouraged from imitating both the Path and the Goal according to the model event, but the Manner could be included free of conflict with other components. Furthermore, Manner was not included in the calculation of the Path bias score. In this phase, Manner was included by two-year-olds a fair amount (mean Manner score .62) and almost all of the time by four-year-olds

(mean Manner score .90). Path bias scores were compared across age groups and task orders using a two-way ANOVA. A main effect of age group was found ($F(1,27) = 18.559$, $p = <.001$), but no main effect of task order or interaction. Four-year-olds were much more Path biased (mean score = .53) in Choice imitation than two-year-olds (mean score = -.02). Next, we compared the Path bias score to chance. Chance in this instance refers to an average Path bias score of 0, having neither a Path nor a Goal preference. In a one sample t-test it was also found that the four-year-old bias was significant ($t(15) = 4.977$, $p > .001$), but the two-year-old bias was not ($t(14) = -.269$, n.s.).This difference is shown in Figure 4.

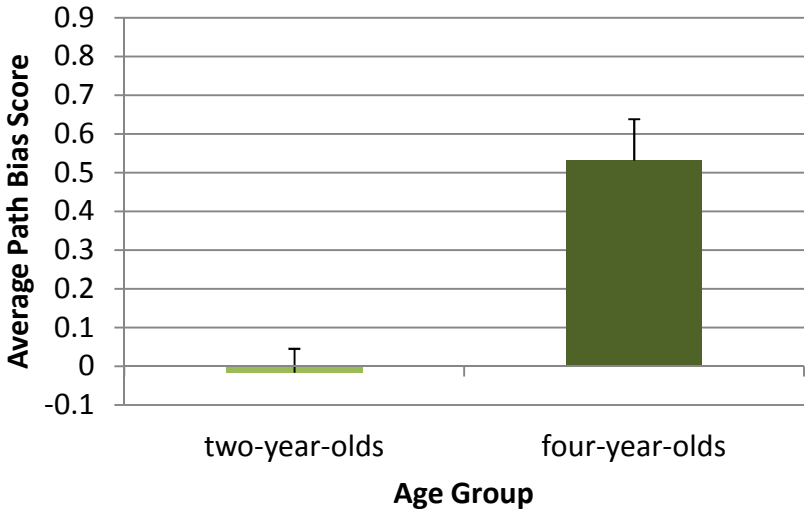

Figure 4: Average Path bias score for participants in each age group, in Choice imitation

Description

Older children included more in their descriptions than younger children, as predicted, and this was supported by three different dependent measures of linguistic complexity in two-way ANOVAs. The factors for this analysis were age group and task order. For the first measure, number of clauses per description, two-year-olds uttered an average of 1.08 clauses,

while four-year-olds averaged 1.77 clauses ($F(1, 27) = 33.732$, $p < .001$). For each description, on average two-year-olds only mentioned 1.20 event components, while four-year-olds mentioned 2.24 ($F(1, 27) = 35.118$), $p < .001$). Even per clause, a stricter measurement, four-year-olds included more components—two-year-olds included an average of 1.02 event components, whereas four-year-olds included 1.36 ($F(1, 27) = 6.825$, $p = .015$). No effects of or interactions with task order were found using these measures, suggesting that order of task did not affect how much was included in participants' descriptions.

To analyze inclusion of event components, a repeated measures ANOVA was conducted with task order (imitation first vs. description first) and age group (two-year-olds vs. four-year-olds) as between subjects independent variables and event component (Path vs. Manner vs. Goal) as a within subjects independent variable. The dependent variable was the proportion of correctly described elements, calculated from the average component inclusion scores from each participant's descriptions. An effect of age group was found ($F(1,27) = 39.692$, $P < .001$), and a three-way interaction between age group, task order, and component inclusion was found ($F(2) = 4.075$, $P = .022$). No other significant main effects or interactions were found. The age group effect reflects the immediately previous finding that four-year-olds include more information in their descriptions than two-year-olds, including more of each event component. As shown in Figure 5, the three-way interaction highlights the greater Path inclusion for two-year-olds in the imitation first condition. To investigate this difference further, paired sample t tests comparing Path to other components in both conditions for the two-year-old age group revealed that Path was included more than Manner ($t(6) = -2.520$, $P = .045$) or Goal ($t(6) = -2.763$, $P = .033$) in the imitation first condition but not the description first condition, and this is shown in Figure 5. As with their imitation scores, this suggests a Local effect for two-year-olds, that those who imitated

first were drawn to the Path of the event, or that those who described first were drawn away from the Path of the event. Four-year-olds do not show this Local effect, however they hit ceiling in their inclusion of all components (average Manner inclusion score = .70, Path = .84, Goal = .81), making it difficult to determine whether they preferred any information in this task.

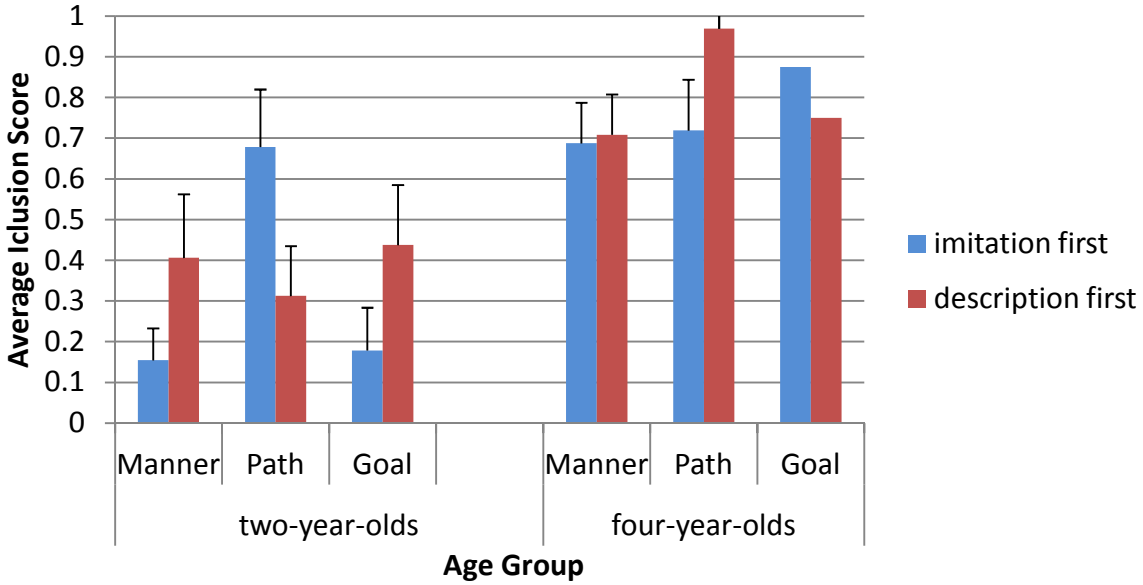

Figure 5: Description inclusion scores by each age group, broken down by task order condition

Syntactic Encoding

In each age group, clear trends about what forms children used to encode specific event components emerged. Table 1 shows the structures used by children in each age group to encode Manner, Path, and Goal, by clauses. Children in both age groups never used an adjunct such as "by hopping" to encode Manner, they always used the main verb ("He slid," 53.07 mos). To encode Path, the overwhelming majority of children used the prepositions "up" and "down" (e.g. "He went up," 51.73 mos), though twice in each age group a Path verb was used ("He climbed on top of the top," 34.43 mos). Finally, for Goal, participants almost always used a separate

clause to describe the Goal event, "He went in the bowl" (32.47 mos), but there was one exception in each age group in which more than just the Goal was described in a clause ("He went up and down and into it," 29.8 mos). These means of encoding events are shown in Table 1.

Description Inclusion

Two-year-olds

Manner		Path		Goal	
Omitted	38	Omitted	26	omitted	37
main verb	16	preposition	26	clause	16
Adjunct	0	Verb	2	physical	1
Total	54	Total	54	total	54

Four-year-olds

Manner		Path		Goal	
Omitted	65	Omitted	57	omitted	62
main verb	46	preposition	52	clause	48
Adjunct	0	Verb	2	physical	1
Total	111	Total	111	total	111

Table 1: Inclusion for event components in description clause data for each age group, and by what form

Task Comparison

To analyze performance across tasks, the best comparison is the Direct imitation inclusion scores to the description inclusion scores, and this is for two reasons. First, the procedures between these two tasks are comparable. In these cases, participants were presented with DM events and asked to represent them as they were, either in a description or in an imitation. They did not have to decide between the importance of modeled components as they were asked to do in Choice imitation. Moreover, the inclusion scores from Direct imitation and

description are on the same scale, 0 to 1, representing either no inclusion or inclusion of a component. Path bias scores are on a different scale, -1 to 1, which represents preference toward either the Goal or the Path of the modeled event. Therefore, we will compare inclusion scores for Direct imitation and description for this analysis.

A repeated measures ANOVA was run using age group (two vs. four) and task order (imitation first vs. description first) as between subjects factors, and task (Direct imitation vs. description) and event component (Manner vs. Path vs. Goal) as within subjects factors, and proportion of inclusion as the dependent variable. A four-way interaction was found between component, task, age group, and task order ($F(2,54) = 3.66$, $p = .032$). This finding is similar to the three-way interaction found in the description task. Two-year-olds in the imitation first condition describe more Path, and this pattern is not reflected by four-year-olds, description first participants, or Direct imitation trials. This suggests that two-year-olds' descriptions were influenced by task order.

Summary

Neither twos nor fours showed preferences toward any event components. Four-year-olds correctly imitated all components more, and two-year-olds performed better in the imitation first condition, suggesting that they become fatigued after description. No Global or Local effects were found for this phase. Two-year-olds were neither Path nor Goal biased, while four-year-olds were strongly Path biased, perhaps demonstrating a Global effect. Though the patterns of encoding events were consistent across age groups, two-year-olds show a Local effect that fours do not. In the imitation first condition, they describe more Path than in the description first condition. Four-year-olds perform similarly in both conditions, but they hit ceiling, obscuring

whether they really did not prefer information or were skilled enough simply to encode all event components in their descriptions.

These findings suggest domain specific patterns that were not captured by our initial hypotheses about language and thought. Younger children were Locally influenced in both tasks by the first task they undertook, while older children did not perseverate. They had already developed their domain-specific ways of analyzing events to carry out these tasks.

Discussion

The aim of this study was to investigate the nature of the relationship between language and thought by comparing older and younger children's analyses of space in a linguistic vs. a non-linguistic domain. We expected to find Global or Local effects from language onto non-linguistic thought. Furthermore, we suggested that if these effects existed, then older children should show different preferences for event components than younger children. Given that we found a Global and a Local effect, it appears as though our third Hypothesis was borne out; however, we expected the same effects to appear in both age groups. Instead, each age group unexpectedly demonstrated a different kind of effect. Older children showed a Global effect – they were Path biased in imitation Choice in either condition whereas younger children were not. However, we expected this preference to be reflected in their descriptions, which it was not. Four-year-olds hit ceiling in the description task, having enough linguistic prowess to describe virtually all components of the event. For Local effects, we expected older children to show a stronger Local influence from language onto thought because of their more sophisticated linguistic resources that they could rely on as a strategy in a non-linguistic task. However, they showed no difference in behavior across conditions in imitation Choice, suggesting no Local

effect. Contrary to expectations, it was the younger participants that showed a Local effect in the imitation first condition. When they imitated first, they included Path more often in their descriptions.

A possible explanation for the Global effect shown only in the older participants' Choice imitations is that gaining language experience does cause a Global effect from language onto thought, such that the older children actually do prefer Path in their descriptions and carry this preference over into their event representations. Taking this view, the older children's linguistic Path preference was hidden by their ceiling effect in the description task, which did not force them to prioritize components in the same way that the forced imitation choice paradigm did. In order to determine whether older children's stronger Path bias develops from a more sophisticated linguistic perspective or simply a better understanding of the framing of the event, a better linguistic bias measure is needed.

A follow-up study of 16 four-year-olds was conducted at the same satellite research space as the current study. The task was modeled closely on the Choice phase of the imitation task, but the children were instead forced to choose between two descriptions of an event. Children were shown a DM event from the same set of clips used in the current description task, and then the experimenter showed them two puppets that each provided a description of the event. One puppet described just the Path of the event, and the other just the Goal, and children were asked, "Who said it better?" Because both descriptions were accurate, if the description encoding one component is chosen more often than the other we can conclude that that component is more important to the child's understanding of the event in a linguistic domain. Also, crucially, children cannot hit ceiling in this task. However, from preliminary analysis we found that the four-year-olds preferred neither the Path nor the Goal description of events, suggesting that they

prefer neither event component. Perhaps this finding can explain our null effect in the four-year-old description task, if they described all components at equal levels, than they may not have preferred any component of the event in their linguistic analysis.

The Local effect observed in the younger but not the older children in this study suggests that language learning really does change the way tasks in different domains are handled. Two-year-olds, not having developed a method (rhetorical style) for analyzing the event when approaching the task, develop an ad hoc strategy from the onset and use it to complete the current and subsequent tasks. This perseverating strategy looks different depending on what the initial domain is. In the context of the study, when participants described the event first, having no rhetorical style to draw from that outlines which event components are more important to encode, they chose components at random and formed their utterances. From the data, we saw that they showed no preference for any of the components and included them roughly equally often. Then, when they performed Choice imitation, they also chose which components to imitate randomly, showing no preference for either Path or Goal. On the other hand, participants who imitated first were drawn to the Path of the event, perhaps because the ramp was the largest and most dynamic piece of the events' setting, and so described more Path. Older children did not need to rely on ad hoc strategies, they already knew how to analyze an event in a particular domain. Their Path preference in Choice imitation remained regardless of the domain they analyzed the event in first.

Due to some methodological limitations of the study that we did not foresee, a potentially informative Global effect from language onto thought was obscured by the ceiling effect observed in our four-year-olds' descriptions. For the younger children, description task was difficult enough that for any given movie they were not skilled enough to encode of all of the

event components. This fact created a certain amount of competition between the event components as any of them could potentially be described by the speaker. This competition made the task similar to the Choice phase of the imitation task, in which the Path and the Goal were directly in conflict with one another. However, the older children did not need to make this choice among components in the description task, as their language skills were sophisticated enough to describe all of the event components at once. The Direct phase of the imitation task was similarly easy for both age groups, although the description first two-year-olds perform worse, perhaps due to fatigue from the difficult description task. Although the description and Direct imitation tasks were designed to create an equal opportunity to represent any of the event components, four-year-olds' inclusion rates were so high that we could not determine whether their similar inclusion rates across tasks were due to the design of the tasks or a roughly equal interest in all of the event components. Future directions of this research will need to account for the skill of the children in the design, perhaps by presenting them with more complex events, or asking them to perform a more difficult task. For example, consider the following hypothetical experiment. Children are shown simple DM events and then are shown an event in which two of the event components have changed and are asked to describe the difference between the two events. If they have prefer one event component over the other in their linguistic analysis of the events, they may describe one component more often than the other.

The relationship between language and thought may be the combination of Global and Local influences from one domain onto the other. As we have seen from the current study, these effects can change as children get older and gain language skill depending on the demands of the task at hand.

References

Bunger, A., Trueswell, J., & Papafragou, A. (2010). Seeing and saying: The relation between event apprehension and utterance formulation in children. Proceedings from the 34th Annual Boston University Conference on Language Development. Somerville, MA: Cascadilla Press.

Choi, S. & Bowerman, M. (1991). Learning to express motion events in English and Korean: The influence of language-specific lexicalization patterns. *Cognition* 41, 83-121.

Cifuentes-Férez, P. & Gentner, D. (2006). Naming motion events in Spanish and English. *Cognitive Linguistics* 17(4), 443-462.

Gennari, S., Sloman, S., Malt, B, & Fitch, W. (2002). Motion events in language and cognition. *Cognition* 83(1), 49-79.

Gleitman, L., January, D., Nappa, R., & Trueswell, J. (2007). On the give and take between event apprehension and utterance formulation. *Journal of Memory and Language* 57(4), 544-569.

Gleitman, L., Papfragou, A. (2005). Language and thought. In Keith J. Holyoak and Robert G. Morrison (ed.), *The Cambridge handbook of thinking and reasoning*, 633-661.

Goldin-Meadow, S. & Zheng, M. (1998). In Carruthers, P. & Boucher, J. (Eds.), *Langauge and thought*, 26-54. Cambridge: Cambridge University Press.

Jackendoff, R. (1990). *Semantic structures*. Cambridge, MA: MIT press.

Johanson, M., Selemis, S., & Papfragou, A. (2009). Cross-linguistic biases in the semantics and acquisition of spatial language. Proceedings from the 33rd Annual Boston University Conference on Language Development. Somerville, MA: Cascadilla Press.

Lakusta, L. & Landau, B. (2005). Starting at the end: The importance of goals in spatial language. *Cognition* 96(1), 1-33.

Levinson, S. (2001). Covariation between spatial language and cognition. In M. Bowerman, & S. Levinson (Eds.), *Language acquisition and conceptual development*, 566-588. Cambridge: Cambridge University Press.

Lucy, A. (1992). *Grammatical categories and cognition: A case study of the linguistic relativity hypothesis*. Cambridge: Cambridge University Press.

Papafragou, A. (2007). Space and the language-cognition interface. In P. Carruthers, S. Laurence & S. Stich (Eds.), *The innate mind: Foundations and the future.* Oxford: Oxford University Press.

Papafragou, A., Massey, C. & Gleitman, L. (2002). Shake, rattle, 'n' roll: The representation of motion in language and cognition. *Cognition* 84(2), 189-219.

Papafragou, A., Massey, C., & Gleitman, L. (2006). When English proposes what Greek presupposes: The cross-linguistic encoding of motion events. *Cognition* 98(3), 75-87.

Papafragou, A. & Selimis, S. (2010). Event categorization and language: A cross-linguistic study of motion. *Language and cognitive processes* 25(2), 224-260.

Pulverman, R., Hirsh-Pasek, K., Golinkoff, R., Pruden, S., Salkind, S. (2006). Conceptual foundations for verb learning: Celebrating the event. In K. Hirsh-Pasek & R. Golinkoff (Eds.), *Action meets word: How children verbs*, 134-159. Oxford: Oxford University Press.

Slobin, D. (1996). From Thought and Language to Thinking for Speaking. In J. Gumperz & S. Levinson (Eds.), *Rethinking linguistic relativity* (70-96). Cambridge: Cambridge University Press.

Slobin, D. (2003). Language and thought online: cognitive consequences of linguistic relativity. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and thought*. Cambridge, MA: MIT Press.

Slobin, D. (2004). The many ways to search for a frog: Linguistic typology and the expression of motion events. In S. Strömqvist & L. Verhoeven (Eds.), *Relating events in narrative, vol. 2: Typological and contextual perspectives* (219-257). Mahwah, NJ: Lawrence Erlbaum Associates.

Slobin, D. (2005). Linguistic representations of motion events: what is signifier and what is signified? In C. Maeder, O. Fischer, & W. Herlofsky (Eds.), *Iconicity inside out: Iconicity in language and literature* 4 (307-322). Amsterdam/Philadelphia: John Benjamins.

Slobin, D. L. (2006). What makes manner of motion salient? Explorations in linguistic typology, discourse, and cognition. In M. Hickmann & S. Robert (Eds.), *Space in Language: Linguistic systems and cognitive categories*, 59-81. Amsterdam/Philadelphia: John Benjamins.

Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. In Timothy Shopen (ed.), *Language typology and syntactic description, vol. 3: Grammatical categories and the lexicon*, 57-149.

Wagner, L., and Lakusta, L. (2009). Using language to navigate the infant mind. *Perspectives on Psychological Science* 4 (2), 177-184.

Wagner, L., A. Yocom, M. & Greene-Havas (2009). Children's understanding of directed motion events in an imitation choice task. *Journal of Experimental Child Psychology* 100, 264-275.

Whorf, B. (1956). In J. B. Carroll (Ed.), *Language, thought and reality: selected writings of Benjamin Lee Whorf*. Cambridge, MA: MIT Press.